



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

***LIBOPT – An environment for testing solvers on
heterogeneous collections of problems
– Version 1.0 –***

J. Charles GILBERT — Xavier JONSSON

N° ????

22 février 2007

_____ Thème NUM _____

 *apport
technique*



LIBOPT – An environment for testing solvers on heterogeneous collections of problems – Version 1.0 –

J. Charles GILBERT* , Xavier JONSSON†

Thème NUM — Systèmes numériques
Projet Estime

Rapport technique n° 0000 — 22 février 2007 — 33 pages

Abstract: The Libopt environment is both a methodology and a set of tools that can be used for testing, comparing, and profiling solvers on problems belonging to various collections. These collections can be heterogeneous in the sense that their problems can have common features that differ from one collection to the other. Libopt brings a unified view on this composite world by offering, for example, the possibility to run any solver on any problem compatible with it, using the same Unix/Linux command. The environment also provides tools for comparing the results obtained by solvers on a specified set of problems. Most of the scripts going with the Libopt environment have been written in Perl.

Key-words: benchmarking – collection of problems – CUTer – Libopt – Modulopt – optimization – performance profile – scientific computing – solver comparison – testing environment.

* INRIA Rocquencourt, projet Estime, BP 105, 78153 Le Chesnay Cedex, France; e-mail: Jean-Charles.Gilbert@inria.fr.

† Mentor Graphics (Ireland) Ltd. - French Branch; 180, Avenue de l'Europe - Zirst Montbonnot; F-38334 Saint Ismier Cedex; e-mail: Xavier.Jonsson@mentorg.com.

LIBOPT – Un environnement pour évaluer des solveurs sur des collections hétérogènes de problèmes

– Version 1.0 –

Résumé : L’environnement Libopt est à la fois une méthodologie et un ensemble d’outils qui peuvent être utilisés pour tester, comparer et profiler des solveurs sur des problèmes de diverses collections. Ces dernières peuvent être hétérogènes dans le sens où leurs problèmes peuvent avoir des caractéristiques communes qui diffèrent d’une collection à l’autre. Libopt apporte un point de vue unificateur sur ce monde composite en offrant, par exemple, la possibilité de lancer n’importe quel solveur sur n’importe quel problème compatible avec lui, en utilisant la même commande Unix/Linux. L’environnement fournit également des outils pour comparer les résultats obtenus par divers solveurs sur un jeu spécifié de problèmes. La plupart des scripts qui accompagnent l’environnement Libopt ont été écrits en Perl.

Mots-clés : calcul scientifique – collection de problèmes – comparaison de solveurs – CUTEr – environnement de test – évaluation de performance – Libopt – Modulopt – optimisation – profil de performance

1 Introduction

Two of the issues that come up with software development have to do with benchmarking and profiling solvers on some (possibly large) collections of problems. For example, in the optimization community, these issues are frequently dealt with the *CUTEr testing environment* [1, 7]. This one supplies ready-to-use interfaces between some known solvers and a collection of problems encoded with the SIF language. It is intended to help developers to test their optimization code. The Libopt package has been designed for a similar purpose but, unlike CUTEr, it provides neither collections of problems written in a specific format, nor decoders for converting problems written in some language into FORTRAN or C; on the other hand, it is not restricted to optimization. Rather, Libopt has been thought up for organizing and using problems coming from heterogeneous sources. Heterogeneity refers, in particular, to the variety of languages used to write the problems. As a result, CUTEr is considered by Libopt as a particular collection of problems having its own features and solver-collection interface.

Originally, we had in mind to make the problems of the *Modulopt collection* [8] as easily available as the CUTEr problems, despite the diversity of their encoding. The Modulopt collection is formed of problems coming from industrial or scientific computing sources. Because of their complexity, these codes are often not modeled on the academic standard of idealized problems, so that it would have been difficult and boring to rewrite them in SIF. This state of affairs motivated the development of the Libopt environment, which resulted in a layer covering various collections of problems (including CUTEr and Modulopt) and solvers, organizing and normalizing the communication between them.

In addition to its solver-collection setting, Libopt also provides a number of programs, mainly Perl scripts, that perform repetitive tasks, such as running solvers on problems, collecting their results, and comparing them. The features of the solvers and collections are actually encoded in these programs, some of them having to be written for each solver-collection pair. Once these are available, all the results are obtained and compared using the same Unix/Linux commands. This is one of the interest of the proposed approach. Now, Libopt is an open structure, which can grow, depending on the needs of its users, by incorporating more solvers, problems, or even collections of problems. A great part of this paper describes this aspect of the software.

The suffix “opt” of the environment name, and of its various companion scripts and programs, reveals the fact that Libopt was first introduced and used to deal with optimization solvers and problems. However, the concepts implemented in this software are sufficiently general for being suitable to other scientific computation fields. The tools have been designed with this idea of generality in mind.

The paper is organized as follows. The guided tour of section 2 invites the reader to discover the main aspects of the software. This view of Libopt, from 10,000 feet, essentially presents the tools that are intended to be routinely used. Section 3 describes the Libopt package and its directory structure; it enumerates the installation instructions. A deeper understanding of the Libopt mechanisms is necessary if one wishes to enrich

the environment with new collections of problems and solvers; these “squalid details” are presented in section 4. We consider in sections 5 and 6 respectively how to install a new collection of problems and a new solver. The paper ends with a prospective section that presents possible improvements.

Notation. We use the following typographic conventions. The `typewriter` font is used for a text that has to be typed literally and for the name of files and directories that exist as such (without making substitutions). In the same circumstances, a generic word, which has to be substituted by an actual word depending on the context, is written in *italic typewriter font*.

Throughout, we assume that the operating system is Unix or Linux. The *command prompt* is denoted by the character ‘%’. Optional arguments in a command line are surrounded by the brackets ‘[’ and ‘]’.

We use the following abbreviations of regular expressions:

`\u` for `[_\t]` (the *blank* character),
`\s` for `[_\f\n\r\t]` (the *space* character),
`\w` for `[a-zA-Z0-9_]` (the character for writing words),

where `\f` stands for a formfeed, `\n` for a newline, `\r` for a carriage return, and `\t` for a tab.

2 A guided tour

The Libopt environment has been designed to make easier and faster the repetitive tasks linked to testing, comparing, and profiling solvers on problems coming from heterogeneous collections of problems. These tasks can be divided into three stages: running solvers on problems, gathering the results, and comparing them. Libopt has three commands (and some others) associated with these stages: `runopt`, `addopt`, and `perfopt`. This section offers an introduction to them. Note that manual pages are available, which give the details on the use of each of these commands. But before this, let us start by defining terms that are continually used in this paper.

2.1 A few definitions

The *Libopt hierarchy* is the set of directories and files that form the Libopt environment. The directory from which the Libopt scripts are launched is called below the *working directory*. The scripts take care that this directory is not in the Libopt hierarchy. If this were the case, there would be a danger of incurable destruction. Indeed, a script like `runopt` generally removes several files from the working directory after a problem has been solved.

A *solver* is a computer program that can find the solution to some classes of problems. Hem! Not sure this is very explicit, but we shall not try to be more precise; in

particular, we shall take the notion of *problem* as a premiss, not requiring any definition. The *name* of a solver is a string that must be match by `\w+` (see the notation above for the meaning of the character ‘`\w`’; the multiplier ‘`+`’ means that one or more of the immediately previous character or character class, here ‘`\w`’, must be present).

A *collection* is a set of problems in an arbitrary scientific computation area. The *name* of a collection is also a string that must be match by `\w+`. Problems in the same collection must differ by their *name* (again a string that must be matched by `\w+`), but two problems belonging to two different collections may have the same name. There are good reasons for gathering problems having common features in the same collection. For instance, problems in a collection are often written in the same language (Fortran, C, Matlab, Scilab, Ampl, Gams, SIF, to mention a few). The motive is that, in the Libopt environment, all the problems of a collection are dealt with the same scripts and that these scripts are easier to conceive if the problems are written in the same language. For the same reason, problems in a collection usually belong to the same scientific computation domain (optimization, linear algebra, differential equations, etc). Another property that is usually shared by all the problems of a collection is the extend to which they are protected against dissemination; this is useful for determining the public to which the collection can be distributed.

A *subcollection* is a subset of problems belonging to the same collection. Its *name* must also be match by `\w+`. A collection may contain several subcollections, and a problem may belong to more than one subcollection. The reason why the notion of subcollection is introduced is clear: some solvers can sometimes only solve a part of the problems of a given collection and it is useful to be able to designate them. For example, in optimization, there is some interest in distinguishing the subcollections of unconstrained problems, of linear problems, of quadratic problems, of bound constrained problems, etc, since there are solvers that are dedicated to these classes of problems.

2.2 Running solvers with runopt

Libopt can only deal with solvers that are registered in its environment. The procedure to do this is described in section 6. Below, we denote by

solv

the generic name of such a solver. Similarly, Libopt can only consider collections of problems that are installed in its hierarchy. This simply means that some directory (or a symbolic link to it) has to contain the problems of the collection (these can be in an arbitrary format) and some files and scripts have to tell how to use the collection (this is why the format can be arbitrary). This subject is discussed in section 5. Below, we denote by

coll, *subc*, and *prob*

the generic names of a collection, subcollection, and problem, respectively.

The simplest way of running the solver *solv* on the problem *prob* of the collection *coll* in the Libopt environment is by entering

```
% echo "solv coll prob" | runopt
```

In other words, the standard input of the script *runopt* is fed with the *command string* “*solv coll prob*”. We quote a few interests in proceeding like this.

- First, the form of the command for running any solver on any collection is invariant: it does not depend on the solver, the collection, or the problem. In our experience, this property saves much memorization effort. In particular, it gives to a collection, which is not used often and is coded in a manner that is difficult to remember, more chance to be tested, even after it has been abandoned for several years.
- Second, it defines a standard, to which solver developers can contribute by providing the interfaces between their solver and various collections.
- Also, the possibility to consider a large diversity of collections should allow the environment to accept problems coming from various sources.

A slightly more powerful use of the *runopt* command is

```
% echo "solv coll.subc" | runopt
```

where *subc* is a subcollection of the collection *coll*. By this command the solver *solv* is run on all the problems of the subcollection *subc*. If the suffix “*.subc*” is not present in the command string above, the *all* or *default* subcollection is assumed, depending on the presence or absence of problems in the command string. Subcollections are described by lists of problems (see section 4.2), which are searched in various directories by *runopt* (see section 4.3).

2.3 Gathering the results with *adopt*

By the *runopt* command a solver-problem pair writes its output on its usual files, which probably include the Unix/Linux standard output. In order to compare solvers, there is no reason to save all these files, which mainly interest the developer of the problem code. In fact, the standard Libopt scripts normally remove these files after having run a solver on a problem (this behavior can be prevented by setting the option *-k* of the *runopt* command, see the introduction of section 4 and section 4.3.2). This is because, Libopt is designed to *compare* the results of various solvers, not to *analyze* them. For this reason, Libopt is interested in the value of various counters that reflect the performance of a solver on a particular problem. In optimization, these counters are often the number of function or derivative evaluations, the CPU time, the precision of the obtained solution, and many others.

In order to be able to make comparisons, Libopt imposes that the results relevant to a comparison be condensed on the standard output in a string of the form

```
libopt%solv%coll%prob%sequence-of-token-number-pairs
```


This one is called the *Libopt line*. It is formed of a sequence of fields separated by the character ‘%’.

- The first field is the string “*libopt*”. It is present to make it easier to locate the Libopt line in the standard output (using the command `grep` for example). Therefore this string must appear only once in the standard output and the Libopt line cannot be split on several lines on the standard output.
- The next three fields give in order the name of the solver (*solv*), the name of the collection (*coll*), and the name of the problem (*prob*), whose relevant results are given in the following fields.

These first four fields are positional, i.e., their order is imposed. This is not the case for the following ones.

- The *sequence-of-token-number-pairs* is a string formed of a sequence of token-number pairs, again separated by the character ‘%’. A *token-number pair* is a string of the form

token=number

The character ‘=’ must separate the *token* (a string matched by `\w+`) from the *number* (a string representing a real number). There must be at least two token-number pairs, one of which is used to compare the results (it must be a *performance* token-number pair actually, see below) and another one must have the form

info=number

where the string “*info*” is imposed and a *number* equal to 0 means that the solver *solv* has successfully solved the problem *prob*.

Note that, since the Libopt line is written by some program provided by the developer of a solver, it is that program that decides whether the solver has successfully solved a given problem; Libopt has no means to take such a decision.

The libopt line can be written with some flexibility: blanks (matched by `\u+`) surrounding the various fields and elements of fields are ignored, and a comment can be introduced (it starts with the sharp character ‘#’ and goes up to the end of the line).

An example of Libopt line in optimization could be

```
libopt%m1qn3%modulopt%u1mt1%n=1875%nfc=143%nga=143%info=0
```

to mean that the solver *m1qn3* has been run on the problem *u1mt1* of the collection *modulopt*, that this problem has 1875 variables (*n=1875*) and that a solution has been found (*info=0*) using 143 function and gradient evaluations (*nfc=143* and *nga=143*).

Now, the command

```
% echo "solv coll.subc" | runopt | grep libopt \  
> solv_coll_subc.lbt
```

gathers in the file `solv_coll_subc.lbt` a sequence of Libopt lines describing the behavior of the solver `solv` on the problems of the subcollection `subc` of the collection `coll`. It is interesting to save this file preciously, since the previous command may have required much time to run, since it is not a large file, and since its ascii encoding makes it very stable with respect to the possible evolutions of the languages (Libopt, Perl, or Unix/Linux).

There may be many files of results like `solv_coll_subc.lbt` and there is some interest in gathering all the results they contain in a single file. This gathering operation is also a good opportunity to verify that the Libopt lines in the result files are consistent and that there is at most one result for each (solver, collection, problem) triple. This is exactly what the command `addopt` does. If one enters

```
% adopt solv_coll_subc.lbt
```

the Libopt lines in the result file are decoded, checked (see below), and stored in a database, whose default name is `dtbopt` in the working directory. This database is no longer an ascii file, but a binary file or a pair of binary files, depending on the operating system. Possible file names are `dtbopt.db` or the pair `dtbopt.dir/dtbopt.pag`. This variety of storages makes the database not very portable, which is also a reason to save the ascii files that have been used to generate it. The database stores a collection of key-value pairs: the key is the first part of the Libopt line, more exactly the string “`solv%coll%prob`”, without useless blanks (there is no reason to store the invariant string “`libopt%`”); the value is the second part of the Libopt line, namely what we have denoted the *sequence-of-token-number-pairs* above. By entering the `addopt` command for all the relevant result files, one can obtain a database containing all the results of interest, without duplicated or contradictory data. The database can be managed, using the options of the `addopt` command: `-r` for replacing entries and `-d` for deleting results. A full description of the `addopt` command line is given in section 4.4 and in its manual page.

The amount of verifications done on the libopt lines by `addopt` depends on the presence and contents of the `~/liboptrc` startup file (see section 4.1 for a complete description of this file). If there is no such file, `addopt` just verifies the conventions mentioned in the beginning of this section. It cannot do more. In particular, it makes no assumption on what are the tokens, on the quantity of token-number pairs, and (clearly) on their order. These pieces of information are actually very domain dependent and certainly not identical in optimization, in linear algebra, or in the differential equation domain. However, to avoid typos, you can list the valid tokens in the startup file `~/tcshrc`. If this file is present, `addopt` will read it, and if it contains the directive `tokens`, `addopt` will verify that the tokens in the Libopt line are among those listed by this directive.

2.4 Comparing the results with perfopt

The `perfopt` command has been designed to make comparisons between solvers on a selected set of problems. The comparison is based on the results stored in a database,

one generated by the `addopt` command. A single criterion can be used for making this comparison, and it must be one of the tokens present in the Libopt lines summarizing the results to compare. This comparison made by `perfopt` produces files that can be subsequently dealt with Matlab or Gnuplot. These files describe the *performance profiles* à la Dolan and Moré [4] of the solvers.

In the Libopt line, one finds descriptive token-number pairs (or descriptive tokens) and performance ones. The semantics behind this distinction is rather intuitive: a *performance token* is a token that can be used to compare solvers, while a *descriptive token* is a non-performance token. A performance token-number pair must have the following properties:

- the token-number pair must, obviously, depend on the solver (not on the problem),
- the number of the token-number pair must be positive (> 0),
- the number of the token-number pair is *better* (i.e., indicates a better performance) when it is *smaller*.

In the example given on page 7, `n=1875` must be considered as a descriptive token-number pair (the dimension of the problem is not a property of the solver), while `nga=143` is normally a performance one, since an optimization solver can be considered to be better when it requires less derivative evaluations. Libopt cannot verify that these rules have been respected (except for the positivity of the number), but the performance profiles that `perfopt` generates assume that they hold; they would have no meaning otherwise. On the other hand, if you want Libopt to make some verifications on the correct use of performance tokens, you can list them in the startup file `~/.tcshrc` (see section 4.1 for more precision on this file).

Note, however, that descriptive tokens may also be interesting for selecting the problems on which a comparison is made. For example, one can be interested in making a comparison of solvers on problems with more (or less) than 1000 variables, and/or those with (or without) inequality constraints, etc. To see how to prescribe this, read what is said about the “problem directive” in the description of the `perfopt.spc` specification file in section 4.5 or in the manual page of `perfopt`.

The easiest way of using the script `perfopt` is by entering

```
% perfopt
```

Then, the script scrutinizes the database (its default name is `dtbopt` in the working directory) and reads the specification file `perfopt.spc` in the working directory to know what it has to do. This file is described in detail in section 4.5. It can contain many *directives*. Three of the most important ones are those that specify the solvers that have to be compared, the problems on which they have to be compared, and the performance token chosen for the comparison. For example

```
solver bosch klee durer
collection painting.selfportraiture
performance time
```

is understood by `perfopt` as a requirement to compare the execution time (if this is the meaning of the token `time` in the third line) for the three solvers `bosch`, `klee`, and `durer` (first line), on the problems from the subcollection `selfportraiture` of the collection `painting` (second line). Then `perfopt` generates files with the default names `perf.gnu` and `perf.m`, which subsequently processed by Gnuplot or Matlab can give a graph like the one in figure 1 (a few additional tunings in the specification file are

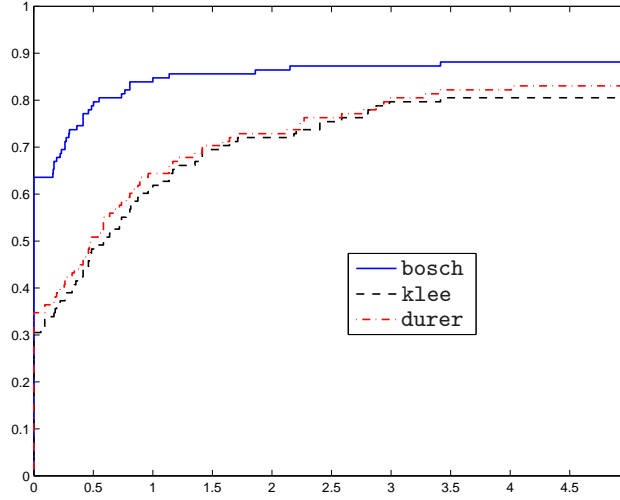


Figure 1: Typical performance profiles for three solvers

necessary to get precisely this graph).

Performance profiles have been introduced by Dolan and Moré [4]. To be comprehensive, we feel it necessary to give a few words on the meaning of these curves. More can be found in the original paper. The main objective is to replace tables of numbers by curves (one per solver), which, with some reading-keys, provide a rapid understanding of the *relative* performance of various solvers on a given set of problems. More specifically, these curves are used to compare the efficiency of a set \mathcal{S} of solvers on a set \mathcal{P} of test problems. Let

$$\tau_{p,s} := \text{performance of the solver } s \text{ on the problem } p,$$

where the *performance* refers to the number of a token-number performance pair. The *relative performance* of a solver s (with respect to the other solvers in \mathcal{S}) on a problem p is the ratio

$$\rho_{p,s} = \frac{\tau_{p,s}}{\min\{\tau_{p,s'} : s' \in \mathcal{S}\}}.$$

Of course $\rho_{p,s} \geq 1$. On the other hand, it is assumed that $\rho_{p,s} \leq \bar{\rho}$ for all problems p and solvers s , which can be ensured only by setting $\rho_{p,s}$ to the large number $\bar{\rho}$ if the solver s cannot solve the problem p (`info` is nonzero in the Libopt line). Actually, we shall consider that s fails to solve p if and only if $\rho_{p,s} = \bar{\rho}$. The *performance profile* of

the solver s (relative to the other solvers) is then the function

$$t \in [1, \bar{\rho}] \mapsto \varphi_s(t) := \frac{|\{p \in \mathcal{P} : \rho_{p,s} \leq t\}|}{|\mathcal{P}|} \in [0, 1],$$

where $|\cdot|$ is used to denote the number of elements of a set (its cardinality).

Only three facts need to be kept in mind to have a good interpretation of these upper-semi-continuous piecewise-constant nondecreasing functions:

- $\varphi_s(1)$ gives the fraction of problems on which the solver s is the best; note that two solvers may have an even score and that all the solvers may fail to solve a given problem, so that it is not guaranteed to have $\sum_{s \in \mathcal{S}} \varphi_s(1) = 1$;
- by definition of $\bar{\rho}$, $\varphi_s(\bar{\rho}) = 1$; on the other hand, for small $\varepsilon > 0$, $\varphi_s(\bar{\rho} - \varepsilon)$ gives the fraction of problems that the solver s can solve; this value is independent of the performance token chosen for the comparison;
- the value $\varphi_s(t)$ may be given an interpretation by inverting the function $t \mapsto \varphi_s(t)$: for the fraction $\varphi_s(t)$ of problems in \mathcal{P} , the performance of the solver s is never worse than t times that of the best solver (this one usually depends on the considered problem); in this respect the argument at which φ_s reaches its “almost maximal” value $\varphi_s(\bar{\rho} - \varepsilon)$ is meaningful.

With performance profiles, the relative efficiency of each solver appears at a glance: the higher is the graph of φ_s , the better is the solver s .

3 The Libopt package

This section contains a brief explanation on how to unpack and install the Libopt package, and a rather detailed discussion of its directory/file structure.

3.1 Retrieving the environment

Libopt has been designed to be used by more than one person at the same time, so that it is recommended to put the Libopt hierarchy in a location that is visible to all the potential users. The Libopt hierarchy can be retrieved by one of the following two manners.

- The first possibility is to get the package from the Inria GForge <http://gforge.inria.fr>, using the `svn` command

```
% svn checkout svn+ssh://name@scm.gforge.inria.fr/svn/libopt
```

where *name* stands for a registered name. This method requires to be registered to the GForge and to have set up the `ssh` communication. It is therefore more appropriate for the developers of (and the contributors to) Libopt.

- The other possibility is to retrieve the tar-gzip file

`libopt.tar.gz`

and to unpack it, using successively the commands `gunzip` and `tar`.

3.2 Structure of the Libopt hierarchy

The retrieving procedure above creates a tree of subdirectories, whose root name is `libopt`. This top-level directory contains the directories `bin` (Perl scripts of the Libopt commands), `collections` (problem collections), `doc` (some documentation), `man` (manual pages), `platforms` (platform descriptions, see section 4.7), and `solvers` (solver descriptions and interfaces); see figure 2.

Level-0	Level-1	Level-2	Description
<code>bin</code>			Libopt scripts
<code>collections</code>			
	+--- <code>cuter</code>		CUTer collection
	+--- <code>modulopt</code>		Modulopt collection
<code>doc</code>			Documentation
<code>man</code>	+--- <code>man1</code>		Manual pages
<code>platforms</code>			Platform descriptions
<code>solvers</code>	+--- <code>m1qn3</code>	+--- <code>bin</code>	M1qn3 binary directory
		+--- <code>cuter</code>	Interface M1qn3/CUTer
		+--- <code>modulopt</code>	Interface M1qn3/Modulopt
	+--- <code>sqppro</code>	+--- <code>bin</code>	Sqppro binary directory
		+--- <code>cuter</code>	Interface Sqppro/CUTer
		+--- <code>modulopt</code>	Interface Sqppro/Modulopt

Figure 2: Part of the Libopt hierarchy in the standard distribution

The `collections` directory gathers a set of subdirectories, each of them corresponding to an installed collection of problems. In the standard distribution, one finds the two collections CUTer [7] in `cuter` and Modulopt [8, 5] in `modulopt`. More generically, the directory `collections/coll` contains the description of the collection `coll`. What this actually means is reflected in the scripts that use this collection, so that these directories can be organized with a great freedom, as far as Libopt is concerned. More is said about this directory in section 5.

It is natural to put in these collection directories (the subdirectories of `collections`) some *lists of problems*. These are files whose name must have the suffix “.1st” (see also section 4.2). Two of these lists are mandatory:

- `all.1st` is the list of all the problems of the collection;

- `default.lst` is a list of a subset of the problems of the collection; this list is chosen by some scripts when no list is specified as one of their arguments (see section 4.3.2); it is sometimes a symbolic link to the list `all.lst`.

It is also natural to have in `collections/coll` a subdirectory, usually called `probs`, which contains all the problems of the collection. Of course `probs` can be a symbolic link to the actual problem directory. For example, in our Libopt hierarchy, `collections/cuter/probs` is a symbolic link to a directory containing the gzipped SIF encoding of the CUTer problems. On the other hand, in the standard distribution, `collections/modulopt/probs` contains a subdirectory for each of the problems of the Modulopt collection, which describes this problem by programs, data, and a makefile.

The directory `collections` provides no information on the solvers. This information is given by the directory `solvers`, which has a subdirectory for each solver known to the Libopt environment: `m1qn3` and `sqqpro` in the standard distribution. `M1qn3` is an unconstrained optimization solver, using the ℓ -BFGS formula [6], and `sqqpro` is a nonlinear optimization solver, using an SQP approach [2, 3]. The directory `solvers` is actually very rich, since, in some sense, it reflects the structure of the Cartesian product

$$\{\text{sol}_1, \dots, \text{sol}_m\} \times \{\text{coll}_1, \dots, \text{coll}_n\} \quad (3.1)$$

corresponding to the interfaces between solvers and collections. Indeed, for each pair $(\text{sol}_i, \text{coll}_j)$, where sol_i is some solver `soli` and coll_j is some collection `collj`, the directory `solvers/soli/collj` contains information describing how to run the solver `soli` on the problems of the collection `collj`. This directory includes the following mandatory files/scripts (see also figure 3):

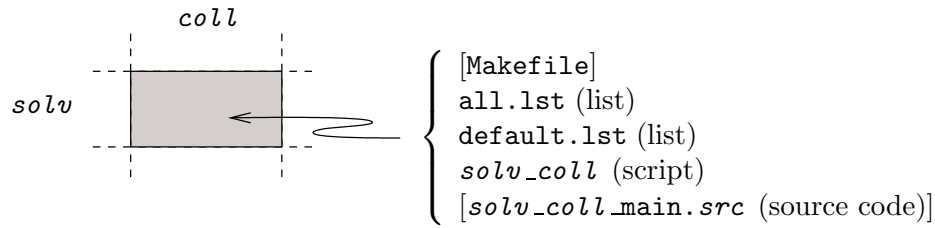


Figure 3: A single cell of the $\{\text{solvers}\} \times \{\text{collections}\}$ Cartesian product

- `all.lst` is the list of all the problems of the collection `coll` that the solver `sol` is structurally able to solve (note that this meaning is quite different from the one of the file `collections/coll/all.lst` described above); for instance, the file `solvers/m1qn3/cuter/all.lst` is a symbolic link to the list of unconstrained problems of the CUTer collection;
- `default.lst` lists a subset of the problems in `all.lst`; this list is used by some scripts when no list is specified as one of its arguments (see section 4.3.2);
- `sol_coll` is a script that tells how to run the solver `sol` on a problem of the collection `coll`; it essentially takes care of the operating system commands that

are required to launch *solv* (see section 4.3.3); note that the name of the script depends on the directory where it is placed.

The directory may also contain other files, like the source code of a main program *solv_coll_main.src* able to run the solver *solv* on a single problem of the collection *coll*, which takes care of the instructions that cannot be written in a Unix/Linux script, and a makefile *Makefile* (see section 4.3.4).

3.3 Installation instructions

Once the Libopt has been unpacked, a few things need to be done before being able to use the environment. We suggest the following five steps.

1. First, add the following environment variable definitions to the shell startup file `~/.tcshrc` (the `tcsh` shell is assumed; adapt the format and use another startup file if the shell is different; echoing your `$SHELL` environment variable might be useful here):

```
setenv LIBOPT_DIR    dir
setenv LIBOPT_PLAT   plat
setenv PATH          ${PATH}:${LIBOPT_DIR}/bin
```

The variable `$LIBOPT_DIR` sets the root directory of the Libopt hierarchy; hence change “*dir*” into the location of this root directory. The variable `LIBOPT_PLAT` provides the *platform* on which you work, see section 4.7 for valid value for the string “*plat*”. The third setting adds to `$PATH` the directory of the Libopt commands. In some Linux system, this setting automatically adds `$LIBOPT_DIR/man` to the possible search paths for the manual pages. If this is not the case with your system, you might have to set

```
setenv MANPATH       ${MANPATH}:${LIBOPT_DIR}/man
```

With these settings you should have an appropriate answer to the following commands:

```
% runopt -h
% man runopt
```

2. You can now launch the installation procedure (from any directory, since this script is in `$LIBOPT_DIR/bin`):

```
% install_libopt
```

This procedure defines the files `collections/collections.lst` (the list of collections existing in the standard distribution), `solvers/solvers.lst` (the list of solvers existing in the standard distribution), and for each solver *solv* in the previous list, it defines `solvers/solv/collections.lst` (the list of collections that the solver *solv* has been prepared to consider). It also verifies the consistency of the Libopt hierarchy, by checking whether expected files are present.

3. This third step is optional but recommended. It consists in introducing the startup file `~/.liboptrc` (hence in your home directory). This file provides additional information to some Libopt commands, for example for helping `addopt` to detect typos in Libopt lines. See section 4.1 for a detailed description of the directives that can be put in this startup file.

In optimization, one can start by copying a file given in the `doc` directory

```
% cp $LIBOPT_DIR/doc/liboptrc_optim ~/.liboptrc
```

and modify it afterwards if the tokens used in the file does not suit to ones needs.

4. The fourth step deals with the installation of a collection. The Modulopt collection is already installed in the Libopt hierarchy, so that nothing need to be done for being able to use it. Very little has to be done for the installation of the CUTER collection: follow the procedure detailed in section 5.1.
5. This fifth step is the most complex. It deals with the installation of a solver in the Libopt environment. Indeed, Libopt does not provide solvers.

If you are lucky, you might have one of the solvers of the standard distribution, those now listed in the file `solvers/solvers.lst`. In this case, the installation is usually rapid; it is described in the file `solvers/solv/README.install` if `solv` is the considered solver. Otherwise, follow the steps given in section 6.

You should now be able to run a solver (say `solv`) on a problem (say `prob`) of some collection (say `coll`), by typing

```
% echo "solv coll prob" | runopt
```

The next section gives more details on the possibilities of the Libopt environment and explains how the commands are dealt with.

4 Libopt in depth

Section 2 has presented the tools the most frequently used of the Libopt environment. Sometimes, however, it is necessary to do other operations like introducing a new collection of problems or a new solver in the environment. A deeper understanding of the Libopt mechanisms is required for realizing these operations safely. The goal of this section is to provide a comprehensive description of the principles governing the software, as well as the available commands.

The Libopt commands are located in the `bin` directory of the hierarchy. They are written in Perl. Many of these commands accept several of the following options, whose meaning will not be repeated:

- h help mode; a short help message describing the command usage is printed;
- k keep mode; the files generated by the command and its children are not removed on exit;
- t test mode; same as -v, but the commands are not executed;

`-v` verbose mode; the commands are executed and are also printed on the standard output.

When a command launches other scripts and some of the options `-k`, `-t`, and/or `-v` are present, these options are transmitted to these scripts.

4.1 The `~/.liboptrc` startup file

The file `~/.liboptrc` is used to provide some additional information to Libopt, in connection with the activity of the user. Here are the *directives* that are meaningful:

```
tokens = list-of-tokens
performance_tokens = list-of-performance-tokens
data_base = DBFile
```

The first two directives are related to the Libopt line (see page 6). The *list-of-tokens* is a blank-separated list of the tokens that are considered to be valid in the Libopt line. If this directive is present, `addopt` uses the tokens of this list to determine whether the tokens encountered in the submitted Libopt lines are valid (otherwise there is no verification). The *list-of-performance-tokens* is a blank-separated list of tokens that can be considered as performance tokens. If the directive `tokens` is present, it is verified that the performance tokens are among the tokens in the *list-of-tokens*. The performance tokens are used by `perfopt` and `queryopt` to determine whether the token proposed to make the comparison of solvers is a performance token.

The third directive, `data_base`, specifies the database name *DBFile* that has to be used when this one is not specified in the command line of some commands that need such a database (`addopt`, `perfopt`, ...). Hence the name given on the command line has priority over the name given in the startup file. If neither the command line nor the startup file provide a file name, the commands assume that the database is the file `dtbopt` in the working directory. The actual name of the database on disk may have the extension `".db"` or may be represented by two files with names having the extensions `".dir"` and `".pag"`; this depends on the operating system.

An example of `liboptrc` file can be found in the directory `doc` (under the name `liboptrc_optim`). It may be useful in optimization.

4.2 Lists

The behavior of the Libopt scripts depends on various lists, which are used to control the consistency of the commands with respect to the content of the Libopt hierarchy. These lists are located in various directories and, for clarity, they have the suffix `".lst"`. One finds lists of solvers, of collections, and of problems (*subc.lst*):

```
$LIBOPT_DIR/collections/collections.lst
$LIBOPT_DIR/collections/coll/subc.lst
$LIBOPT_DIR/solvers/solvers.lst
```

```
$LIBOPT_DIR/solvers/solv/collections.lst
$LIBOPT_DIR/solvers/solv/coll/subc.lst
```

A list is simply a sequence of strings separated by spaces (`\s+`). *Comments* are possible; they start from the character ‘#’ up to the end of the line. Comments are used to describe the list and its elements.

4.3 The runopt command

4.3.1 Overview

The `runopt` command is probably the most complex of those offered by the Libopt environment. It is used to run solvers on problems (see section 2.2 for an introduction). Because of its genericity, this command cannot realize itself all the details of this operation. Indeed, it cannot know all the features of any possible solver and any possible collection of problems (even those that are still not installed in the hierarchy!). For this reason, `runopt` must delegate part of the operations to other scripts/programs, some of them being written by the persons introducing new solvers and collections into the Libopt environment. The operation of running solvers on problems is actually decomposed in three levels of programs.

- The Libopt level.

The first level is formed of the operations implemented in the `runopt` script itself. These operations are those that are independent of any solver and any problem collection. The aim of this level is to analyze the `runopt` command line and to decompose it in a sequence of elementary operations of the form

run the solver *solv* on the problem *prob* of the collection *coll*. (4.1)

There are as many elementary operations as there are combinations of solvers, collections, and problems expressed by the `runopt` command lines. To execute each elementary operation, `runopt` calls the program *solv_coll* of the second level.

The `runopt` script is further described in section 4.3.2 below. Note that it has not to be modified when the Libopt environment is enriched with new solvers or new collections of problems.

- The operating-system level.

The second level is the one that takes in charge the elementary operation (4.1). It is realized by the programs

```
$LIBOPT_DIR/solvers/solv/coll/solv_coll
```

These are actually Perl scripts in the standard distribution. There is a third level since, by choice and for more flexibility and readability, the *solv_coll* programs are restricted to the necessary operations *at the operating system level*. However,

unlike `runopt`, these operations depend here on the solver `solu` and the problem collection `coll`. Typical operations consist in copying data files in the working directory, running makefiles to get executable programs, launching the main program, and removing data and result files from the working directory on exit from the main program. For some collections of problems (like CUTer), some of these operations may also have been delegated to other programs intimately incorporated into the collection.

More is said about the `solu_coll` programs in section 4.3.3. Note that such a program has to be written when a new solver or a new collection of problems is installed in the Libopt environment. It can often be adapted from an existing one; for example all the `solu_cuter` programs only differ by a few character strings.

- The main program level.

Running a program on a problem does not only depend on operations done at the operating system level. A main program has to be written, which reads the data, calls the solver, and delivers diagnosis. It is at this level that the *Libopt line* is generally written on the standard output. This is the part of the Libopt environment that is the most interlinked with the solver and problem structures. As a result, such a program has to be written for each new solver-collection pair. This level is further described in section 4.3.4.

4.3.2 The runopt script

The general form of the `runopt` command is the following

```
% runopt [-h] [-k] [-t] [-v] [CommandFile]
```

where *CommandFile* is a file that contains command lines of the form

```
solu[.tag] coll[.subc] [list-of-problems]
```

in which *solu* is some solver, *tag* is an optional string tagging the solver name in the Libopt line (the usefulness of this option is discussed in section 4.6), *coll* is some collection, *subc* is an optional subcollection of *coll*, and *list-of-problems* is an optional blank-separated list of problems of the subcollection. If there is no file of commands, `runopt` reads the standard input. Therefore, if there is just a single command line, it may be easier to feed the standard input of `runopt`, using the Unix/Linux command `echo` like on page 6.

For each command line, the list of problems that are actually tried to be solved by *solu* is established by `runopt` in the following manner.

1. If there is no subcollection specified in the command line (no “*.subc*”), `runopt` assumes either the “*all*” subcollection if the *list-of-problems* is nonempty or the “*default*” subcollection otherwise. The logic behind this rule is to be non-restrictive when some problems are mentioned in the command line (the “*all*” subcollection is supposed to give all the problems of the collection), and to limit the number of

problems to consider when no problem is specified (the “default” subcollection is supposed to give a not too large subset of typical problems of a huge collection).

2. Once a subcollection, say *subc2*, has been determined, either from the one given in the command line or from the rule just mentioned, **runopt** looks for a file describing the subcollection *subc2*. It searches in the following order:

```
coll.subc2.lst in the working directory
$LIBOPT_DIR/solvers/solv/coll/subc2.lst
$LIBOPT_DIR/collections/coll/subc2.lst
```

The logic is the following. Since **runopt** first looks in the working directory, the user is allowed to build a temporary list of problems for a special purpose. Next, priority is given to the lists the solver *solv* has declared in connection with the collection *coll*, since it is the solver that knows the type of problems it can solve. If no file *subc2.lst* is found in the above locations, the command line is ignored.

3. **Runopt** takes the additional precaution of eliminating from *subc2* the problems that are not soluble by *solv* (i.e., those that are not in *\$LIBOPT_DIR/solvers/solv/coll/all.lst*). Let *subc3* be the resulting list.
4. The final list is then, either *subc3* if there is no *list-of-problems* in the command line, or the intersection of *subc3* and the *list-of-problems* if the latter is nonempty. Let *subc4* be the final list.

Then **runopt** launches the script *\$LIBOPT_DIR/solvers/solv/coll/solv_coll* for each problem *prob* of the list *subc4*.

4.3.3 The *solv_coll* scripts

A *solv_coll* script must exist for each solver *solv* that has been installed in the Libopt environment to run problems from the collection *coll*. It must be placed in the directory *\$LIBOPT_DIR/solvers/solv/coll* and is launched by the **runopt** command. It must accept the following command line structure

```
% solv_coll [-k] [-t] [-v] prob
```

The options *-k*, *-t*, and *-v* are inherited from the **runopt** command that launches *solv_coll*. It is the **runopt** command that determines the problem *prob* to solve (see section 4.3.2). As discussed in section 4.3.1, the *solv_coll* program has in charge the operations, at the operating system level, for running the solver *solv* on the problem *prob* of the collection *coll*.

When there is already a solver, say *solv0*, that has been installed to run problems from the collection *coll*, it is often quite easy to get the correct *solv_coll* script. It is indeed usually sufficient to copy

```
% cp $LIBOPT_DIR/solvers/solv0/coll/solv0_coll\
    $LIBOPT_DIR/solvers/solv/coll/solv_coll
```

and to substitute in the copied file every instance of the string “*solv0*” by the string “*solv*”. Examples exist in the standard distribution, which have this property due to a high degree of parametrization of the scripts.

Consider now the general case, when one has to start from scratch, because a new collection *coll* is installed in the hierarchy. Let us describe the operations that the script *solv_coll* must realize, for a particular solver *solv*. The problem to solve, *prob*, is given by the single plain argument of the command. We have said above that the script has to take care of the options *-k*, *-t*, and *-v* of the command. Next, the script *solv_coll* must realize three tasks.

1. Build in the working directory the executable program, say *solv_coll_main* (a binary code that will solve the problem *prob* using the solver *solv*), and copy into the working directory, from somewhere in the hierarchy `$LIBOPT_DIR/collections/coll`, the data files that are useful for running *solv_coll_main*.

These operations depend in depth on the installation of the collection *coll*.

- For the Modulopt collection, the operations are encoded at appropriate targets in the makefiles:

```
$LIBOPT_DIR/collections/modulopt/probs/prob/Makefile,  
$LIBOPT_DIR/solvers/solv/modulopt/Makefile.
```

Therefore, *solv_coll* just execute these targets after having set the two environment variables used by them, which specify the problem name (`$MODULOPT_PROB`) and the working directory (`$WORKING_DIR`).

- For the CUTER collection, this first task could have been skipped, since the CUTER command *sdsolv* (or something similar) takes it in charge. However, for allowing to store the problems in a compressed format, the file *prob.SIF.gz* is copied in the working directory and decompressed there.
2. Run the executable program *solv_coll_main* (or *sdsolv* for the CUTER collection) in the working directory.
 3. Remove from the working directory the now useless executable program *solv_coll_main* (not necessary for the CUTER collection), the data files, and the output files that have been generated during the execution of the program. This task also depends on the collection *coll* and the solver *solv*. For the CUTER collection, part of the task is taken in charge by the command *sdsolv* itself.

4.3.4 The main program

This is the main program that would have to be written if one has to run the solver *solv* on a particular problem *prob* of the collection *coll*. The required genericity of this main program (it has to be able to consider *any* problem *prob* of the collection) is obtained in the following way. All the problems of the collection are described by

procedures with a name that depends on the realized function but *is identical for all the problems*. Typically, one finds

- *initialization procedures*, which specify the dimensions of the problem, possibly its name, initialize the variables, read the data, etc,
- a *simulator*, which gives the state of the system to solve at a given iterate when the problem is nonlinear and the solver has an iterative nature,
- *auxiliary procedures*, which precise the way some objects are computed (for example, the procedure that realizes the inner product used for computing a gradient),
- and possibly *post-solution procedures*, which can realize some post-solution computations.

The connection between the main program and the problem is then realized by the link editor, to which the object files associated with the main program and the procedures describing the selected problem are given by the *solu_coll* script described in section 4.3.3.

It is normally the main program that prints the *Libopt line* on the standard output. It can do this after the problem has been solved by collecting the various features of the problem and the various counters that depict the behavior of the solver *solu* on the problem *prob*.

4.4 The adopt command

We have discussed in large in section 2.3, the role of the **adopt** command in the basic work cycle of the Libopt environment. In particular, all the details on the Libopt line have been presented in that section. In the present section, we make it precise the form of the **adopt** command line.

The accepted forms of the **adopt** command are the following

```
% adopt [-h] [-t] [-v] [-b DBFile] [-r] ResFile
% adopt [-h] [-t] [-v] [-b DBFile] -d selection
```

If the option **-h** is present, **adopt** prints a short help message describing the command usage and exits. The first command is used for adding results in a database, the second for deleting results.

The database *DBFile* is specified with the **-b** option. If this option is not present, **adopt** looks for the database specified by the directive **data_base** in the startup file *~/.liboptrc*. Finally, if this one does not exist, **adopt** assumes that the database is the file *dtbopt* in the working directory (see section 4.1 for the details). The file *ResFile* used in the first form of the command is supposed to contain Libopt lines, typically generated by the **runopt** command. A result corresponding to a triple (solver, collection, problem), already existing in the database is not replaced, unless the option **-r** is present.

The string **selection** used in the second form is interpreted as a file name if it is not terminated by the character **'%'** or as a way of selecting triples (solver, collection,

problem) otherwise. If `selection` is view as a file name, this file is supposed to be formed of Libopt lines and the results with the same triples (solver, collection, problem) as in the Libopt lines are deleted from the database. If `selection` is the string “`solv%coll%prob%`”, the result corresponding to the triple (solver, collection, problem) = (`solv`, `coll`, `prob`) is deleted from the database (if present). An empty solver (resp. collection, problem) field in the string `selection` matches any solver (resp. collection, problem). For instance, if `selection` is the string “`solv%%`” or simply “`solv%`”, all the results corresponding to the solver `solv` are deleted from the database. As another example, if `selection` is the string “`%coll%prob%`”, all the results corresponding to the problem `prob` of the collection `coll` are deleted from the database. The option `-t` can be used to see the effect that a deleting command would have before really doing it.

4.5 The perfopt command

The `perfopt` command was introduced in section 2.4, where its role in the basic work cycle of the Libopt environment was shown. In this section, we describe the command in detail. It has the following form:

```
% perfopt [-h] [-v] [-b DBFile] [-p ptok] [-log] [-g GFile]
```

The script is used to compare the performance of various solvers on a given set of problems. A single criterion is used for the comparison and it can be specified by the option “`-p ptok`”, where `ptok` is one of the *performance tokens* in the Libopt line (see sections 2.3 and 4.1); another way of specifying this performance criterion is to use the specification file `perfopt.spc` (see below). The results on which the considered solvers are compared are supposed to be stored in the database `DBFile` (if the `-b` option is not used, `perfopt` uses the name given in the startup file or the default name `dtbopt`, see section 4.1 for the details), typically built by the `addopt` command. The `perfopt` command generates two files: `GFile.m` and `GFile.gnu` (the names `perf.m` and `perf.gnu` are used if the option `-b` is not present in the command line). The file `GFile.m` is a Matlab function providing a graphic representation of the performance profiles of the compared solvers, `GFile.gnu` is a Gnuplot data file of the same performance profiles. The option `-log` requires that the performance profiles must have a logarithmic x-coordinate.

The `perfopt` command is best tuned by using *directives* in the specification file `perfopt.spc` in the working directory. When starting, `perfopt` tries to read this file and terminates if it cannot find and read it. Therefore, *this file is mandatory*. The file is described in the manual page of `perfopt`. To make the discussion that follows understandable, some of the directives of the specification file have to be clarified. First, at least two solvers have to be selected for the comparison, say `solv1` and `solv2`, which is done by one or more directives of the form

```
solver solv1 solv2 ...
```


This is because **perfopt** cannot choose default solvers and it makes no sense to compare a solver with itself. Optionally, some problems may have been selected by one or more directives of the form

```
collection coll[.subc] ...
```

The list of problems specified by this directive is searched in order in the following files

```
coll.subc.lst in the working directory
$LIBOPT_DIR/collections/coll/subc.lst
```

where *subc* is set to the string ‘all’ if it is not specified by the directive. Finally, a performance token must have been specified, say *ptok*, either in the command line of **perfopt** or by a directive of the form

```
performance ptok
```

Here is how **perfopt** selects the actual problems that will be used for the comparison of the solvers specified by the ‘**solver**’ directives. These problems are those that satisfy the following conditions:

- they are among those problems that have been specified by a ‘**collection**’ directive, if any such directive has been used in the specification file **perfopt.spc**,
- they are present in the database *DBFile* with a *ptok* performance token,
- they have not been discarded by the ‘**problem**’ directive (see the manual page of **perfopt**),
- they have been solved by all the solvers specified by the ‘**solver**’ directives.

4.6 Getting results from a modified version of an installed solver

Suppose that some parameters of some solver, say **goya**, are modified and that it is desirable to see the effect of these modifications on the performances of the solver. This situation also occurs during the development of a solver, when it is desirable to see whether a newly developed technique has a good effect, by comparing the results of the new version with the previous ones. We see three ways of realizing this and consider them from the hardest to the easiest one.

1. A first possibility would be to define a full new hierarchy under the **\$LIBOPT_DIR/solvers** directory with the new version of the solver. This results in defining a new row in the $\{\text{solvers}\} \times \{\text{collections}\}$ Cartesian product. This is certainly safe, but it is not an easy and rapid task. In addition, the solver with the new technique may not exist for a long time. Therefore, this option is probably not the best idea.
2. A better way of doing is to modify temporarily the code that generates the Libopt line, so that instead of generating a line of the form

```
libopt%goya.quatro-de-junio% ...
```

where `goya` has been changed into `goya.quatro-de-junio`. The modified Libopt lines can also be obtained by editing the file containing the Libopt lines of the new version of the solver. The `addopt` command will believe that this line has been generated by the solver `goya.quatro-de-junio`, even though judiciously this one does not exist in the Libopt hierarchy. Therefore, the comparison will be made without difficulty with other solvers whose results have been stored in some database (using the command `addopt`), possibly with `goya.tres-de-mayo`, the presumably best version of `goya` obtained so far.

3. Alternatively, instead of modifying the code generating the Libopt lines, one could modify these lines after they have been written in a file, using a text editor. Libopt offers an automatic way of getting the same effect. It is indeed possible to add a *tag* to the solver name in the Libopt line. This is obtained by entering

```
% echo "goya.quatro-de-junio coll prob" | runopt
```

Then, `runopt` concatenates the string `".quatro-de-junio"` to the solver name in the Libopt line of the standard output, which then becomes

```
libopt%goya.quatro-de-junio%coll%prob%...
```

Of course, `runopt` considers that the solver is `goya` not `goya.quatro-de-junio` (because of the dot `'.'`, the latter does not exist in the Libopt hierarchy). This effect is obtained by filtering the standard output of `runopt` with the stream editor `sed`, which only modifies the Libopt lines.

For example, if it is desired to compare the scalar and diagonal running modes of the ℓ -BFGS code `m1qn3` [6] on 143 unconstrained CUTer problems, one can proceed as follows. Run `m1qn3` with its default setting (diagonal running mode), tagging the solver with the string `"diag"` to remember the option used in the run

```
% echo "m1qn3.diag cuter.unc" | runopt | grep libopt \
> m1qn3.diag-cuter.unc.lbt
```

Set the option `imode(1) = 1` in the main program `m1qn3ma.f` used by the CUTer hierarchy to put `m1qn3` in scalar running mode, generate the object file `m1qn3ma.o`, put it in the appropriate place in the CUTer hierarchy, and run again:

```
% echo "m1qn3.scal cuter.unc" | runopt | grep libopt \
> m1qn3.scal-cuter.unc.lbt
```

Add the results obtained in this way in the database:

```
% adopt m1qn3.diag-cuter.unc.lbt
% adopt m1qn3.scal-cuter.unc.lbt
```

Adapt the solver directive of the specification file `perfopt.spc`:

```
solver m1qn3.diag m1qn3.scal
```

Now enter

```
% perfopt
```

This last command generates the file “`perf.m`”, a Matlab M-file, which can be used to plot the performance profiles of the two versions of `m1qn3`, those that are given in figure 4. An accustomed reader can deduced from these curves that the diagonal mode

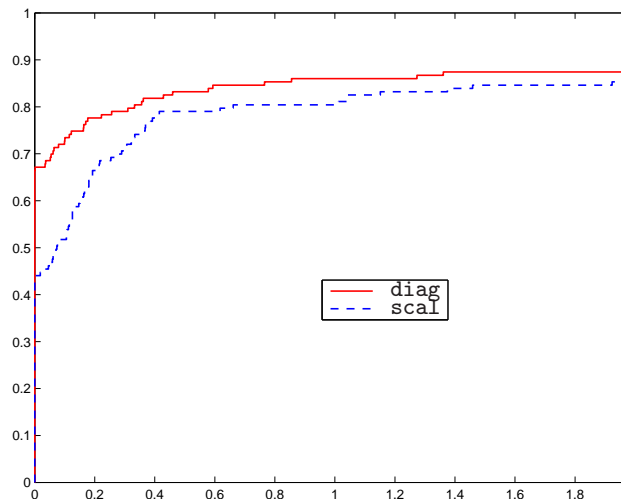


Figure 4: Performance profiles of the diagonal (`diag`, red solid curve) and scalar (`scal`, blue dashed curve) running modes of `m1qn3` on 143 unconstrained problems of the CUTEr collection, comparing the number of function evaluations

of `m1qn3` can be considered as slightly more efficient than the scalar mode (its profile is higher, see section 2.4). This is an average estimation, since the plot with its 2-logarithmic abscissa reveals that, if the scalar mode can require $2^{1.9} \simeq 3.7$ times more function evaluations than the diagonal mode, the latter can also require sometimes $2^{1.4} \simeq 2.6$ times more evaluations than the former.

4.7 Managing platforms

The environment variable `$LIBOPT_PLAT` provides the *platform* on which the user works. This is a string of the form “*mach.os.comp*”, where *mach* designates the machine, *os* is its operating system, and *comp* is the compiler suite. The current standard distribution of Libopt includes the following platforms:

```
mac.osx.gcc
pc.lnx.pg
```

where

```
mac  Macintosh computer (Apple),
```

pc PC-like computer,
lnx Linux operating system,
osx OSX operating system (Macintosh),
gcc GNU Compiler Collection,
pg Portland Group compilers.

The platform name is used by some scripts and makefiles of the Libopt environment. The Libopt hierarchy of the standard distribution contains indeed no object/compiled files. When such files are required they are generated in the working directory and the value of `$LIBOPT_PLAT` is used to generate them. This description is given by files of the form

`make.plat`

where “*plat*” is one of the strings above. These files are located in the directory

`$LIBOPT_DIR/platforms`

A new platform is therefore obtained by adding a new file `make.plat` in that directory, adapting the various variables defined in the file. It is used by resetting the environment variable `$LIBOPT_PLAT`.

5 Hooking a collection of problems onto Libopt

There are probably many ways of hooking a collection of problems to the Libopt environment. This installation depends in great part on the nature of the considered collection. For example, the installation of the CUTer and Modulopt collections are quite different. Almost everything is possible provided the scripts using this collection reflect the options taken during the installation. This is why Libopt can just be viewed as an empty shell (the directory structure) and a methodology (embodied in its scripts).

The installation of a collection is usually quite simple. Two typical installations are those of the CUTer and Modulopt collections. We consider them in sequence.

5.1 Hooking the CUTer collection

Libopt is a layer covering CUTer. Therefore, CUTer must already be installed before being hooked onto Libopt. On the other hand, the standard distribution of Libopt provides part of the installation of CUTer into Libopt, since the directory `$LIBOPT_DIR/collections/cuter` exists and contains some lists of problems `*.lst`. Actually, the only thing that the installed scripts using CUTer require to be done is to define

`$LIBOPT_DIR/collections/cuter/probs`

as a symbolic link to the directory containing the problems of the CUTer collection (files with the suffix “.SIF”), usually denoted by the environment variable `$MASTSIF`.

The scripts present in the standard distribution of Libopt assume that these files are stored in uncompressed or **gzip**-compressed format.

The installation of the CUTER collection is therefore quite simple. This does not mean that it can be readily used. Indeed, the solvers installed in Libopt have now to be prepared to run CUTER problems. This has already been achieved for the solvers in the standard distribution, but some adjustments may be necessary for a new solver; this is discussed in section 6.2.

5.2 Hooking a new collection

In this section, we discuss the installation in the Libopt environment of a typical collection of problems. Let us call it

toys

The installation is modeled on the one of the Modulopt collection, already present in the standard distribution of Libopt [5]. We consider first the definition of the collection into Libopt and next a possible way of introducing problems into the collection.

5.2.1 Defining a new collection in the Libopt environment

1. Define the new directories

```
$LIBOPT_DIR/collections/toys
$LIBOPT_DIR/collections/toys/probs
```

and add the name **toys** to the list

```
$LIBOPT_DIR/collections/collections.lst
```

Put all the problems of the collection **toys** in **probs** (alternatively, you can make **probs** a symbolic link to the directory containing all the problems of the collection). Each problem can be stored in a dedicated directory within **probs**.

2. Create the files

```
$LIBOPT_DIR/collections/toys/all.lst
$LIBOPT_DIR/collections/toys/default.lst
```

The file **all.lst** lists all the problems of the collection **toys**. The file **default.lst** lists a subset of the problems of the collection; this list is chosen by some scripts when no list is specified as one of their arguments; therefore, it can be a symbolic link to the list **all.lst**, obtained using the Unix/Linux command ‘**ln -s all.lst default.lst**’; however, if the collection contains many problems, it is probably better the put in **default.lst** a not too long list of problems reflecting the collection. Other lists of problems can also be defined in the directory **toys**.

All these files may be empty is the collection is started from scratch.

5.2.2 Introducing a new problem in the toys collection

We assume below that the each problem of the `toys` collection is described in the dedicated directory

`$LIBOPT_DIR/collections/toys/probs/prob`.

Hence, let *prob* be the name of the new problem we want to introduce. This can be done as follows.

1. Insert the name *prob* in the list of `toys` problems

`$LIBOPT_DIR/collections/toys/all.lst`

(and possibly in other lists in the same directory). This should be an `ascii` file and *comments* should be possible (they should start from the character ‘#’ up to the end of the line).

2. Create the directory

`$LIBOPT_DIR/collections/toys/probs/prob`,

and put in that directory, all the files that define the problem *prob*: source files, header files (if appropriate), and data files (if appropriate).

3. Create the makefile

`$LIBOPT_DIR/collections/toys/probs/prob/Makefile`,

with the following two targets:

- *prob*, which specifies how to obtain in the working directory an archive *prob.a* with all the object files defining *prob* and which copies into the working directory the data files needed for solving the problem (making symbolic links is probably faster);
- *prob_clean*, which specifies which files has to be removed from the working directory after having solved *prob*.

6 Installing a solver into Libopt

We have seen in section 5 that the installation of a new collection in Libopt is quite simple. However, just as it is, no solver installed in the Libopt environment can use the new collection `toys`, since the cells of the column `solvers`×`{toys}` of the Cartesian product (3.1) have not been filled in (here `solvers` denotes the set of installed solvers, while `{toys}` denotes the singleton made of the single collection `toys`). We consider the following possibilities.

- The solver does not exist in the Libopt environment: then follow the instructions in section 6.1 and consider the other possibilities next.

- The solver exists in the Libopt environment and it is desired that it solves problems from the CUTer collection: see section 6.2.
- The solver exists in the Libopt environment and it is desired that it solves problems from an installed collection different from CUTer: see section 6.3, where the Modulopt collection is considered as a generic case. If the collection is not yet installed, consider first section 5.

Throughout the section, we assume that

baby

is the name of the considered solver.

6.1 Defining a new solver in the Libopt environment

We gather in this section the operations that must be done for defining the new solver **baby** in the Libopt environment, independently of any specific collection. Here they are.

1. The name **baby** has to be added to the file

```
$LIBOPT_DIR/solvers/solvers.lst,
```

which contains the list of solvers of the Libopt environment. This file is considered by Libopt as list of strings separated by ‘\s+’.

2. The following directories must be created

```
$LIBOPT_DIR/solvers/baby
$LIBOPT_DIR/solvers/baby/bin.
```

The directory **baby** is the one that will contain the description of all the collections to which the solver **baby** will have access. The directory **bin** is a possible location for the binaries and libraries defining the solver (or symbolic links to these).

3. Create also an empty file

```
$LIBOPT_DIR/solvers/baby/collections.lst,
```

which contains the list of collections that the solver **baby** can consider (none if this is the first time **baby** is introduced in the Libopt environment).

4. It is recommended to write a file, named `README_install`, that explains how the scripts that you are going to introduce latter in the Libopt hierarchy, under the directory

```
$LIBOPT_DIR/solvers/baby
```

assume to have access to the libraries and binaries of the solver.

That is all for the installation of the solver **baby** itself. Now, this solver cannot access any collection. For this, follow the instructions in section 6.2 for connecting the solver **baby** to the CUTer collection or in section 6.3 for another type of collection.

6.2 Interfacing a solver with the CUTeR collection

We suppose here that

- the solver `baby` has already been installed in the CUTeR environment (for more information, see the manual [7]). This implies, in particular, that a driver or main program `babyma` has been inserted in the CUTeR environment and that the command `sdbaby` can be used, as in

```
% sdbaby prob,
```

where `prob` is the name of a problem in the CUTeR collection.

- the CUTeR collection has been installed in the Libopt environment (see section 5.1),
- the solver `baby` has been introduced in the Libopt environment (see section 6.1).

We are now ready to establish the connection between the solver `baby` and the CUTeR collection, within Libopt. We recommend the following steps.

1. Create the directory

```
$LIBOPT_DIR/solvers/baby/cuter
```

and add the name ‘`cuter`’ to the list of collections that `baby` can solve, namely

```
$LIBOPT_DIR/solvers/baby/collections.lst.
```

2. Define the script

```
$LIBOPT_DIR/solvers/baby/cuter/baby_cuter.
```

The easiest way of doing this is to copy and modify the corresponding script of another solver, say `solv`:

```
% cp $LIBOPT_DIR/solvers/solv/cuter/solv_cuter
    $LIBOPT_DIR/solvers/baby/cuter/baby_cuter
```

The modifications to bring to the content of the file `baby_cuter` are quite simple: change all the instances of `solv` into `baby` (you may want to be careful with some capitalized letters in comment lines). An adaptation of the very last lines of the file `baby_cuter`, dealing with the removal of the files generated by the solver `baby` may require some personal Perl writing (you can just forget this aspect of the script if you prefer, nothing dangerous will occur).

3. Modify the executable main program ‘`babyma`’ installed in the CUTeR environment, so that it prints the Libopt line, defined in section 2.3 (it is also defined in the manual page ‘`libopt`’).

You should now be able to launch your solver `baby` on a CUTeR problem, using the command of Libopt. By typing

```
% echo "baby cuter prob" | runopt
```

you should have the same result as by typing “`sdbaby prob`”.

6.3 Interfacing a solver with another collection

We consider here the example of the interfacing of the solver `baby` with the collection `toys`, defined in the Libopt environment in section 5.2. We recall that the collection `toys` is modeled on the Modulopt collection. More details are given in [5].

1. Create the directory

```
$LIBOPT_DIR/solvers/baby/toys,
```

which will contain the programs/scripts to run the `baby` solver on the `toys` problems, and add the name `toys` to the list

```
$LIBOPT_DIR/solvers/baby/collections.lst,
```

which indicates that `baby` can deal with the `toys` collection.

2. Create the files

```
$LIBOPT_DIR/solvers/baby/toys/all.lst
```

```
$LIBOPT_DIR/solvers/baby/toys/default.lst.
```

- The first file (`all.lst`) must list the problems from the `toys` collection that `baby` is able to solve or, more precisely, those for which it has been conceived. It can contain *comments*, which start with the `#` character and go up to the end of the line. The easiest way of doing this is to start with a copy of the file

```
$LIBOPT_DIR/collections/toys/all.lst,
```

which lists all the `toys` problems, and to remove from the copied file those problems that do not have the structure expected by `baby`. For example, if `baby` is a solver of unconstrained optimization problems, remove from the copied file `all.lst`, all the problems with constraints. Note that other lists might exist in the directory `$LIBOPT_DIR/collections/toys`, which might be more appropriate to start with than the list `all.lst`.

- The second file above (`default.lst`) can contain any subset of the problems listed in the first file (`all.lst`). This file is used as the default subcollection when no list is specified in the `runopt` script. Therefore, it is often a symbolic link to the first file `all.lst`, obtained using the Unix command

```
ln -s all.lst default.lst
```

in the directory `$LIBOPT_DIR/collections/toys`.

3. Create the executable file

```
$LIBOPT_DIR/solvers/baby/toys/baby_toys.
```

This is the script launched by `runopt` to run `baby` on a single `toys` problem. See section 4.3.3 for more information on how to write this script.

4. Create the main program

```
$LIBOPT_DIR/solvers/baby/toys/baby_toys_main.src.
```

This program is very solver dependent and is, with the next step to which it is linked, the most difficult task to realize. It is the main program that will be linked with the subroutines describing the problem from the `toys` collection selected by the `runopt` script, those in the archive `prob.a` (if the selected problem is `prob`) created by the makefile `$LIBOPT_DIR/collections/toys/probs/prob/Makefile` in the working directory (see section 5.2.2).

If Fortran 90/95 is the adopted language, the easiest way to proceed is to copy and rename the file

```
$LIBOPT_DIR/solvers/sqppro/modulopt/sqppro_modulopt_main.f90
```

into the file `$LIBOPT_DIR/solvers/baby/toys/baby_toys_main.f90`. Since this main program is very solver dependent, its part dealing with the solver will have to be thoroughly modified. See [5] for information on the structure of the program.

5. Create the makefile

```
$LIBOPT_DIR/solvers/baby/toys/Makefile.
```

The aim of this makefile is to tell the Libopt environment how to link the solver binary with the object files describing the `toys` problem selected by the `runopt` script. If the latter is `prob`, the corresponding object files will be at link time in the working directory in the archive `prob.a` (see section 5.2.2). The easiest way of doing this is to start with an existing makefile, like

```
$LIBOPT_DIR/solvers/sqppro/modulopt/Makefile.
```

This one will be copied and renamed into the file `$LIBOPT_DIR/solvers/baby/toys/Makefile`.

You should now be able to launch the command

```
echo "baby toys prob" | runopt -v
```

where the flag `-v` (verbose) is used to get detailed comments from the Libopt scripts, which then tell what they actually do. The flag `-t` (test mode) can be used instead, if you want to see what the scripts would do without asking them to do it.

7 Final words

We have presented in this paper the version 1.0 of the Libopt environment and have motivated its main features. This version of the software should certainly deserve improvement: we think, for instance, to the notion of platform, which is presently rather embryonic, to the possible parametrization of problems, and to the development of a Web interface with the environment.

Only very few collections and solvers are incorporated in the standard distribution of this version. Installing new collections and solvers in a consistent way, maintaining the compatibility with the possible evolution of the software, may require some expertise and willpower. Note, indeed, that a change in the software may affect all the cells of the $\{\text{solvers}\} \times \{\text{collections}\}$ Cartesian product. We have the intention to provide these new installations in the future versions of the software, at least in the optimization domain, taking profit of the possible contributions of users.

Acknowledgment

We thank Jorge Nocedal for having offered the second author hospitality at Northwestern University (Evanston, USA) when this work was initiated.

References

- [1] I. Bongartz, A.R.Conn, N.I.M. Gould, Ph.L. Toint (1995). CUTE: Constrained and unconstrained testing environment. *ACM Transactions on Mathematical Software*, 21, 123–160.
- [2] F. Delbos, J.Ch. Gilbert (2005). Global linear convergence of an augmented Lagrangian algorithm for solving convex quadratic optimization problems. *Journal of Convex Analysis*, 12, 45–69.
- [3] F. Delbos, J.Ch. Gilbert, R. Glowinski, D. Sinoquet (2006). Constrained optimization in seismic reflection tomography: a Gauss-Newton augmented Lagrangian approach. *Geophysical Journal International*, 164, 670–684.
- [4] E.D. Dolan, J.J. Moré (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91, 201–213.
- [5] J.Ch. Gilbert (2007). Organization of the Modulopt collection of optimization problems in the Libopt environment. Technical Report 329, INRIA, BP 105, 78153 Le Chesnay, France.
- [6] J.Ch. Gilbert, C. Lemaréchal (1989). Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, 45, 407–435.
- [7] N. Gould, D. Orban, Ph.L. Toint (2003). CUTer (and SifDec), a Constrained and Unconstrained Testing Environment, revisited. *ACM Transactions on Mathematical Software*, 29, 373–394.
- [8] C. Lemaréchal (1980). Using a Modulopt minimization code. Unpublished technical Note.



Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-0803